

A2

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12Q 1/68, C12N 15/11, C07H 21/00,</b> <b>A61K 31/70</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 98/15651</b> <b>(43) International Publication Date:</b> 16 April 1998 (16.04.98)
<b>(21) International Application Number:</b> PCT/GB97/02722 <b>(22) International Filing Date:</b> 6 October 1997 (06.10.97) <b>(30) Priority Data:</b> 9620749.3 4 October 1996 (04.10.96) GB <b>(71) Applicant (for all designated States except US):</b> BRAX GENOMICS LIMITED [GB/GB]; 13 Station Road, Cambridge CB1 2JB (GB). <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> SCHMIDT, Gunter [DE/GB]; Houghton Manor, Houghton, Cambs PE17 2BQ (GB). THOMPSON, Andrew, Hugin [GB/GB]; 25 Knoll Park, Alloway, Ayr KA7 4RH (GB). <b>(74) Agents:</b> DANIELS, Jeffrey, Nicholas et al.; Page White & Farrer, 54 Doughty Street, London WC1N 2LS (GB).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> IDENTIFYING ANTISENSE OLIGONUCLEOTIDE BINDING  <b>(57) Abstract</b>  A method for identifying an antisense oligonucleotide capable of binding to a target mRNA, which comprises contacting the target mRNA with each member of an oligonucleotide library separately under hybridisation conditions, removing unhybridised material and determining which member or members hybridise; wherein the oligonucleotide library comprises a plurality of distinct nucleotide sequences of a predetermined common length, and wherein each nucleotide sequence comprises a known sequence of 4 to 8 bases and all possible combinations of the known sequence are present in the library.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**IDENTIFYING ANTISENSE OLIGONUCLEOTIDE BINDING****Field of the Invention**

The present invention relates to a method for identifying accessible sites in an RNA molecule for antisense agents and to methods for identifying antisense oligonucleotides.

**Background to the Invention**

Most library strategies used for identifying oligonucleotides (ONs) that bind an RNA have relied on cleavage of the target mRNA by RNase H as the key component of their selection process. This, however, only selects for ONs which bind their targets with normal Watson-Crick interactions. ONs that bind their target RNA can mediate cleavage by RNase H and the resultant fragments can be isolated by high resolution gel electrophoresis. The precise sequence of the successful ON is determined by sequencing the fragments. Mishra and Toulmé (1) have developed an alternative selection procedure based on selective amplification of ODNs (Oligodeoxynucleotides) that bind their target and have demonstrated non-Watson-Crick interactions in that binding. Their protocol requires just as much sequencing as the RNase H strategy but they sequence the ODNs and not the target RNA fragments. Furthermore most library strategies use libraries that are free in solution which has the problem of cross hybridisation of ONs within the library. This can be solved by using multiple libraries of minimally cross hybridising subsets but this adds to the labour involved in the selection of ODNs with active antisense properties.

**Summary of the Invention**

The present invention provides a method for identifying an antisense oligonucleotide capable of binding to a target mRNA, which comprises contacting the target mRNA with each member of an oligonucleotide library separately under hybridisation conditions, removing unhybridised material and determining which member or members hybridise; wherein the oligonucleotide library comprises a plurality of distinct nucleotide sequences of a predetermined common length, and wherein each nucleotide sequence comprises a known sequence of 4 to 8 bases and all possible

combinations of the known sequence are present in the library.

Each member of the oligonucleotide library may be contacted with the target mRNA in a separate container which may be immobilised in the container. Alternatively, each member of the oligonucleotide library may be immobilised at a separate location on a hybridisation array. Preferably the length of the known sequence is from 4 to 6 bases.

Each member of the oligonucleotide library may comprise a set of nucleotide sequences. Each nucleotide sequence may have a window region comprising the known sequence and a flanking region of no more than 8 bases, wherein all possible combinations of bases in the flanking region are present in each set and the common length of the nucleotide sequences is no more than 12 bases, preferably no more than 10 bases. The nucleotide sequences may be made from DNA analogues.

In one arrangement, the step of determining which member or members hybridises comprises determining the member or members which hybridise more rapidly.

The sequence of each member which hybridises may be compared with sequence information relating to the target mRNA so as to identify in the target mRNA one or more antisense binding sites.

#### **Description of the Preferred Embodiments**

This invention is aimed at the problem of identifying accessible sites within an mRNA but the object of this technique is to find accessible sites in a system that is applicable to any RNA molecule, that requires no sequencing and is not biased by the use of RNase H. The emphasis of this patent is on the use of libraries of short oligonucleotide (ON) probes, preferably 4-mers. This invention provides a method for determining the accessible sites of an RNA to short ON probes by following the hybridisation reactions of each of the probes in the library with the target. Probes are spatially isolated and in the reactions,

- 3 -

and either the probe or the target RNA is immobilised.

The rationale behind using probes with windows as short as 4mers is the assumption that most occurrences of a given 4-mer will be sequestered in secondary or tertiary structure. The data gained from following each individual hybridisation is interpreted in terms of the primary sequence:- hybridisation of several 4 mers to an accessible region will be marked by clustering of overlapping 4-mers in the primary sequence of the RNA. Ambiguities should be resolvable in part by analysis of the binding kinetics, so even if a 4-mer occurs more than once, it should still be possible in most cases to identify the most accessible regions in the RNA. The importance of using short probes is the ability to calibrate and normalise the hybridisation behaviour of all members of the library, since the library is not so large that this is impractical. A 4-mer library has 256 members, which is a modest number and it is realistic to test such a library against a number of model molecules to calibrate the system. Calibration is important as most 4-mers have different binding energies for a given degree of accessibility. In order to identify the regions of a molecule that are accessible on the basis of hybridisation kinetics requires that one be able to compare normalised kinetic data that takes into account that discrete overlapping 4-mers will bind the same accessible site with different kinetics. Small arrays of probes are cheaper to construct and manipulate when spatial isolation of probes is desired. Furthermore short probes are much less likely to have any secondary or tertiary structure themselves which would complicate the analysis of the hybridisation reactions.

Hybridisation times can be varied to derive kinetic data about each hybridisation reaction. The amount of probe ON hybridised after a given time will give quantitative data about the binding affinity of that probe for its target. By varying the hybridisation duration, a time course for the hybridisation reaction of each probe can be derived. The accessibility of any

- 4 -

region of the structure will be inversely related to the time its complementary probe will take to bind hence detailed information about secondary structure can be acquired by this approach. In combination with a Scintillation Proximity Assay (Amersham), discussed below, the system is still more effective in that no washing is needed, so shorter probes can be used and the hybridisation reaction of each probe can be followed in real time.

**Molecular Probes of RNA Structure:**

Spatial isolation and immobilisation of probes avoids cross-hybridisation problems associated with the use of oligonucleotide libraries free in solution simplifying interpretation of results. This process does not necessarily determine the most effective cut sites for RNase H but provides detailed structural information about the target RNA to allow rational design of targetted molecules which can be then be developed into effective antisense agents, which will not necessarily operate on the basis of RNase H mediated RNA degradation.

**Advantages Over Previous Library Strategies:**

This invention uses random libraries of ONs since, in a random library of ONs of a given length, every possible sequence of that length is represented, hence every sub-sequence of that length that comprises a given target RNA is also represented. This means that a random library will be applicable to any RNA and will contain all the target subsequences within it. Since ONs are monitored independently, this means that this approach entails the targeted libraries discussed in a previous patent PCT/GB96/02275, and gives detailed kinetic behaviour about them. Explicit data concerning accessibility of sub-regions of the RNA to normal Watson-Crick interactions can thus be derived by this system.

Non-Watson-Crick interactions might well be less common than the normal complementary interactions but will also be represented

- 5 -

in a system like this. One would expect interactions with regions that are partially complementary to the target RNA. If Watson-Crick interactions are the only potential interactions that are possible, such partial hybridisation would be expected to be weak but if non-Watson-Crick interactions play a role, as Mishra and Toulmé suggest (1), then subsets of these interactions might be expected to have higher strength interactions with the target RNA. Hence this strategy will identify potentially specific interactions that are non-Watson-Crick.

There appears to be a difference in the results of experiments performed with random libraries and those with targeted libraries. Different cut sites were identified by each library in the experiments performed in our previous patent application PCT/GB96/02275. In those experiments the targeted libraries covered only three specific sub-regions of the target RNA, TNFa. However the sites identified by this library were not identified by the random library. This is probably due to the fact that targeted libraries have fewer members, each of which are at higher relative concentrations than might be the case for members of a random library. The random library, therefore, picked up more accessible sites in regions not covered by the targeted libraries but the ONs corresponding to those used in the targeted library did not pick up the same cut sites. The cut sites picked up by the targeted libraries may have been due to their relatively higher concentration so these may not have been the most accessible. The problem of cross-hybridisation is also difficult to assess in these systems and the contribution of this effect is likely to be different for each type of library. This novel approach should normalise all of these effects giving much more meaningful data.

Kinetic data should also be acquired by this approach, as discussed below, which will give detailed structural information about target RNAs.

#### **Kinetic Data From Hybridisation Experiments:**

- 6 -

Hybridisation times can be varied to derive kinetic data about each hybridisation reaction. The amount of probe ON hybridised after a given time will give quantitative data about the binding affinity of that probe for its target. By varying the hybridisation duration, a time course for the hybridisation reaction can be derived. The accessibility of any region of the structure will be inversely related to the time its complementary probe will take to bind to a given degree hence detailed information about secondary structure should be acquired by this approach.

This procedure would be impossible if libraries of large ONs were used due to the number of ONs that would be required. Calibration of each of the probes in a large library would be a massive task, although with automated image analysis of combinatorially synthesised chips, it might be feasible. However it is not necessary or desirable to use large oligonucleotides.

Real time kinetic assays:

Radiolabelling is a favourable labelling scheme with desirable properties, particularly radio-isotopes that produce low energy radiation with short path lengths such as  $^{33}\text{P}$ . Radiolabelling permits the development of proximity assays. The radiation emitted can be detected by various means, including scintillation or by geiger counters. Proximity detection systems and corresponding proximity assays measure the intensity of a signal from a label which gives a measure of the distance of the label from the detector. Scintillation proximity assays for example are based on the detection of radiation emitted from a radio-isotope. The amount of radiation reaching a scintillant surface is detected by photo-amplification of the scintillation. The mean path length of certain forms of radiation is fairly short. Beta radiation from  $^{33}\text{P}$  has a relatively low energy and short path length. This means a probe labelled with  $^{33}\text{P}$  will be detected only when it is relatively close to the detection surface. The further from the source the lower the intensity that is measured.



If an RNA molecule is labelled with  $^{32}\text{P}$ , and hybridised to an oligonucleotide probe immobilised on a scintillant containing surface or visa versa, one would expect the scintillation count to increase as the amount of RNA hybridised to the probe increases. One would expect there to be a background count, from molecules free in solution close enough to the scintillant surface to be detected, so control reactions with labelled RNA or probe depending on which is to be labelled in the actual assays must be performed. In this sort of system, spatial resolution of probes is essential, as only one radiolabel can be used at a time, so only one probe or RNA could be labelled at a time. The primary benefit of a system based on radiolabelling is the ability to measure hybridisation reactions in real time to give detailed kinetic data. Real time analysis can be achieved by detecting scintillation with a photoamplification and detection system coupled to appropriate signal processing electronics such as the Amersham Cytostar-T scintilling microplates.

#### Hybridisation Probes:

There are two complementary difficulties with the use of short ON libraries. To ensure that the hybridisation of probe to an immobilised RNA is strong enough, one requires that the oligonucleotides be long enough to ensure a reasonable degree of hybridisation but the longer the probe the more massive the task of resolving the behaviour of individual library members. Since RNA/DNA hybrids have higher binding energies than RNA/RNA or DNA/DNA duplexes they are more stable than either of the homo-duplexes, one can use shorter probes than one might use for homogenous interactions and one can use non-natural analogues that have higher binding affinity for natural nucleic acids than probes composed of natural nucleic acids.

#### Non-Natural Nucleic Acid Analogues:

Since this approach does not require recognition by RNase H, one

has more flexibility in the use of non-natural backbones. One can use non-natural base analogues to increase binding energies of any interactions. For the purposes of this system one might use a backbone that is less charged than the natural phosphodiester linkages such as methylphosphonodiester linkages or peptide nucleic acids. Increased energies of probe binding interactions might be useful to allow strand invasion of double stranded regions more easily by the immobilised probed oligonucleotides, the kinetics of which should reveal useful information about the structure of the region being invaded.

Increasing 'Bound' Time of ON Probes:

The problem of weak hybridisation that faces this embodiment is quite acute in the light of the fact that the shorter the ON, the more readily the molecule will dissociate from its target RNA. In carrying out structural probes using short ONs, a large quantity of the probe oligonucleotide would be added to each well of immobilised RNA. This would drive the hybridisation equilibrium in favour of hybridisation, thus significant hybridisation might occur. Once however the unbound oligonucleotide is washed away, the equilibrium shifts dramatically in favour of dissociation of hybridised probe. Various measures can be taken to increase the 'on-time' of bound oligonucleotides by increasing the binding energy of the interactions. Similar considerations about non-natural nucleic acid analogs as discussed in the first embodiment above apply here. One might conceivably also include cross-linking effectors at this stage that are photo-activatable to ensure that hybridised probe is 'fixed'.

Using libraries of short ON probes is a problem in that to get decent hybridisation one needs a nucleic acid of reasonable length, but each additional base added to a probe increases the number of ONs in a library exponentially. One can overcome the problem of weak hybridisation of short oligonucleotides indirectly by constructing oligonucleotides in sets such that

- 9 -

each set is composed of a fixed number of bases of known sequence flanked on one side or other or both by a further fixed number of bases where all possible combinations of nucleotides are represented, these could also be a universal base:

**5' -NNXXXXNN-3'**

The above example has 2 bases, labeled N for any base, flanking the known sequence 'window', XXXX, on both sides. Thus there would be 256 different sets of 8mers which could be used to probe the immobilised RNA. This is effectively the same as probing with 4-mers but the flanking bases would increase the stability of interactions. In conjunction with non-natural nucleic acid analogs, this alternative could increase the stability of complexes sufficiently to allow washing and quantitative measurements to be made.

Using probes with windows as short as 4mers might be problematic in that they may appear too frequently in an RNA and thus might make resolving kinetic data more difficult, as there are likely to be more than 2 occurrences of any given 4mer present in a single RNA. Using 5-mers or 6-mers as windows, with stabilising flanking regions, might be easier to resolve, if automated systems were available to cope with the added labour as these are more likely to appear only once in any given RNA.

#### **Probe arrays:**

One can test oligonucleotides (ONs) individually by spatially isolating probes in separate wells on a microwell plate. One can immobilise a target RNA and challenge the RNA with individual, fluorescently labeled ONs. Thus for a library of ONs of 4 nt, an array of 256 wells would be required, into which equal quantities of RNA would be immobilised. Each well would be challenged with a different labeled ON. The ON is allowed to hybridise for a predetermined length of time and the unhybridised ONs can then be washed off. The quantity of ON hybridised is determined by

measuring the fluorescence in each well. This approach will require significant quantities of RNA which can readily be generated using, for example, the T7 phage RNA polymerase system. An alternative is to immobilise the ON probes and challenge these with labeled RNA.

A random oligonucleotide library can be constructed while immobilised on a glass surface such that distinct regions of the array carry distinct oligonucleotides within the library (2). A random library of ONs of 8 nt will have  $4^8$  possible members if all possible sequences are represented. Since the ONs are all immobilised on the array there will be no cross-hybridisation.

An RNA for which the tertiary structure is unknown can be cloned and produced in quantity in vitro using for example the T7 phage RNA polymerase system. The RNA can be labeled with a fluorescent label. This can then be used to challenge the immobilised library under conditions that favour the adoption of the normal tertiary structure of the RNA, i.e. in vivo cytoplasmic conditions. The RNA is allowed to hybridise onto the array for a predetermined duration and then unhybridised RNAs are washed off. Where hybridisation has occurred can be visualised by detecting fluorescence. The regions of the array from which fluorescence is detected will reveal which ONs the RNA will hybridise to.

These approaches face the problem of weak hybridisation of short ON probe; the minimum size of probe ON will be determined by the strength of interaction necessary to immobilise a large molecule of RNA to the array and allow the RNA to resist being washed off. This would have to be tested empirically but the problem of rapid dissociation should be less severe, when the RNA is the mobile element, as the size of the molecule ought to make it somewhat more sluggish than the small ONs used in the first embodiment. The same approach of creating ONs with short windows of known sequence and flanking regions of random sequence can be readily applied to this embodiment and similar considerations regarding non-natural nucleic acid analogues apply too. Using nuclease

resistant analogues would have the additional advantage of being more reusable than phosphodiester linkages, as one can never completely ensure that there is no contamination by nucleases, etc. which would damage a DNA array.

This embodiment is quantitative, as well, as long as each RNA molecule bears a single fluorescent tag. Thus the regions of the array which give the most fluorescence will be those to which the RNA will hybridise to the most strongly. By repeatedly challenging the array with the target RNA for varying but predetermined durations, one would be able to get detailed kinetic information about the hybridisation reaction occurring at each point on the array. This will reveal regions whose structure must be disrupted to allow hybridisation as these will have much slower kinetics.

An alternative to the use of fluorescent labels would be the use of mass labels cleavably linked to the RNA used to probe the array. Mass labels that are derivatives of photo-excitable compounds of the kind discussed in GB 9700746.2, such as nicotinic, sinapinic or cinnamic acid, can be photocleaved and excited into the gas phase by application of appropriate frequencies of light, ideally using a laser. Such labels could be incorporated into an RNA molecule using a terminal transferase reaction for example or an end ligation. If an RNA molecule labelled with laser excitable mass labels of this kind were hybridised to an array of oligonucleotide probes, the degree of hybridisation of the RNA to distinct regions of the array, corresponding to individual oligonucleotides, could be determined by MALDI (Matrix Assisted Laser Desorption Ionisation) mass spectrometry. Simpler, non-excitable mass labels could also be used. In this embodiment the RNA would be hybridised to an array, and following hybridisation, the supernatant would be removed and the array would be embedded in an excitable agent, of the kinds described above, and ionisation of cleaved labels could be mediated indirectly by excitation of the matrix.

**Rational Drug Design with the Process:****Tertiary Structure Determination:**

Secondary structure modelling of nucleic acids is well developed and in conjunction with data from a system like this it should be possible to develop decent models of tertiary structure of a target RNA. This will start to make so-called 'rational' drug design a much more quantitative process. In conjunction with other methods such as NMR, complete structure determination will probably be possible. At the very least, accurate secondary structure prediction should be achieved. Most likely only minimal further information from other methods would be needed to determine a structure for any given RNA.

**Databases and Drug Selection:**

This system will, with use, provide a comprehensive database of information about RNA tertiary structure which will lead to better theoretical models of RNA structure and allow a much more specifically targeted approach to designing effective antisense agents. The evidence of Wagner et al (3) suggests that short oligonucleotides can show specificity for target RNAs. The qualification that must be made to this suggestion is that short ONs would have to be highly structure specific. With a comprehensive database of RNA structures a system like this might actually pinpoint potential drugs directly once it is established what structural features are widespread and which are rare. Directly searching for rare features will thus be possible making choice of drug candidates much simpler.

**More General Antisense Targets:**

This system is not specific to detecting RNase H sensitive sites and could be used for targeting more general antisense agents at various RNAs in cells. Thus this system has potential for more general applications in developing therapeutic agents and

- 13 -

research tools for targeting more general functional RNAs in vivo such as ribosomes, splicing apparatus and some of the active RNAs that have recently been implicated in sex determination.

Automation:

Both embodiments can be implemented with automated liquid handling systems as the procedure is simple and repetitive. The hybridisation array approach would probably be cheaper in terms of reagents but may be difficult to implement if the binding strength of the interaction between array and target RNA is too small to permit washing of the array. The first embodiment will have a much larger requirement for reagents, although micro-arrays of wells can be constructed to reduce this requirement.

**References:**

- (1) R.K.Mishra, J.J.Toulmé, C.R.Acad. Sci. Paris, Life Sciences 317, 977 - 982, 1994.
- (2) W. Bains, Chemistry in Britain, February 1995, 122 - 125.
- (3) R.W. Wagner et al, Nature Biotechnology 14, 840 - 844, 1996.
- (4) Siew Peng Ho et al, Nucleic Acids Research 24, 1901 - 1907.

**Example**

**Hybridisation of 4-mer oligonucleotides to 2 TNF $\alpha$  cDNA subfragments:**

*Subcloning of the TNF $\alpha$  cDNA clone:*

Clone number 53007 (pAW731), containing the full length cDNA sequence for TNF $\alpha$  in the vector pFC54.t, was bought from the American Type Culture Collection (ATCC), Rockville, Maryland, USA. The freeze dried bacteria were resuspended in 2ml of LB broth and 100ul was spread onto LB agar plate containing 100ug/ml of ampicillin. The plates were then incubated at 30°C overnight. A single colony from the plate was then used to inoculate 100ml

- 14 -

of LB broth containing 100ug/ml of ampicillin and this was incubated at 30°C, shaking at 225 rpm overnight. Following this the bacterial culture was pelleted by spinning the solution at 5300rpm in a Heraeus 17RS centrifuge for 15 minutes. The plasmid was then extracted using the Qiaprep Midi Plasmid Purification Kit (Qiagen Ltd.) according to the manufacturers instructions. 2ul of plasmid was then digested with 10 units of BamH1 and Hind III (Boehringer Mannheim) and 2ul of plasmid digested with 10 units of Hind III in 20ul at 37°C for 2 hours. Digestion with BamH1 and Hind III releases the TNFa insert ( ATCC). The digested plasmids were then run on a 1% agarose gel. The Hind III digest released a band approximately 700bp in size and the BamH1 and Hind III digested released a band approximately 700bp and a band approximately 800bp in size. The 700bp fragment was released by Hind III, corresponds to the 3' end of the TNFa insert, and the 800bp fragment released by Hind III and BamH1, which corresponds to the 5' end of the TNFa insert. These two bands were then cut out of the gel with sterile scalpel blades and spun through GenElute Agarose Spin Columns (Supelco, USA) according to the manufacturers instructions to extract the DNA. The DNA was the precipitated with 2.5 volumes of 100% ethanol and 0.1 volumes of 3M sodium acetate. The solution was then incubated at -20°C for 30 minutes and then spun at 13000rpm in a Heraeus A13 benchtop centrifuge for 15 minutes to precipitate the DNA. The supernatant was then poured off and the pellets washed with 100ul of 70% ethanol and allowed to air dry. The dry pellets were then resuspended in 30ul of water. 5ug of plasmid pBluescript SK(+) (Stratagene) was digested with 10 units of BamH1 and Hind III and 5ug digested with Hind III (Boehringer Mannheim) in 50ul at 37°C for 2 hours and then purified with Chromospin-100 columns (Clonetech) according to the manufacturers instructions. The DNA concentration of the purified bands and digested plasmid was measured by reading the absorption at 260nm and 280nm in a Pharmacia Genequant spectrophotometer. The 700 and 800bp bands were then ligated into the appropriately digested pBluescript using 400 units of T4 DNA ligase (New England Biolabs) for the 800bp fragment and the Rapid Ligation Kit



- 15 -

(Boehringer Mannheim) for the 700bp fragment. 2ul of each ligation was transformed into 50ul of XL-1 BLUE competent cells (Stratagene) by incubating the DNA with the cells for 30 minutes on ice, followed by a 45 second 42°C heat shock and another 2 minutes on ice.

Following this 450ul of SOC medium was added and the cells incubated at 37°C, shaking at 250 rpm, for 1 hour. 100ul of the transformation was then plated onto agar plates containing 100ug/ml ampicillin, 0.1mM IPTG and 40ug/ml X-GAL. The plates were then incubated at 37°C overnight.

White colonies were picked and used to inoculate 2.5ml of LB broth containing 100ug/ml ampicillin and incubated at 37°C, shaking at 250 rpm, for 7 hours. 1ml of the bacterial cultures were then spun into eppendorf tubes and the plasmids extracted with the Qiaprep Spin Plasmid Kit (Qiagen). The presence of the correct fragment was checked by digesting 20% of each minipreped plasmid with the appropriate enzyme(s) and the running on an agarose gel. The remaining bacterial cultures from a plasmid containing the 700bp fragment and from a plasmid containing the 800bp fragment were then each used to inoculate 250ml of LB broth containing 100ug/ml ampicillin which was then incubated at 37°C, shaking at 250 rpm, overnight. Following this the bacterial culture were pelleted by spinning the solutions at 5300rpm in a Hereaus 17RS centrifuge for 15 minutes. The plasmids were then extracted using the Qiaprep Midi Plasmid Purification Kit (Qiagen Ltd.) according to the manufacturers instructions.

#### *In Vitro Transcription*

25ug of each plasmid was linearised by digesting them with 50 units of Cla I in 50ul of 1xbuffer H (Boehringer Mannheim) for 3 hours at 37°C.

The digested plasmids were then purified with Chromospin-100 columns (Clontech) according to manufacturers instructions.

- 16 -

40ul of each plasmid was then used to set up a 200ul in vitro transcription reaction using the RiboMax Large Scale T3 RNA production system (Promega) according to manufacturer's instructions. The remaining 10ul of each plasmid was then used to set up a 50ul in vitro transcription reaction using the RiboMax Large Scale T3 RNA production system (Promega) according to manufacturer's instructions but this time 0.25ul of Fluorescein-12-UTP (Boehringer Mannheim) was included in to each reaction to fluorescently label the RNA. All in vitro transcription reactions were incubated at 37°C for 3 hours.

*Demonstration that RNA can be captured via a Biotinylated Oligo*

A 3' biotinylated oligo (cap 53B), with sequence 5'CCACTAGTTCTAGAGCCGCCGCCACCGCGG3' biotin, which would be complementary to the 5' terminus of the RNA produced from the T3 RNA promoter of pBluescript SK (+) and a 5' biotinylated oligo (cap 35B), with sequence 5' biotin-CCCCCCTCGAGGTCGACGGTATCGATAAG3' were immobilised to the wells of a black, streptavidin coated microtitre plate (Boehringer Mannheim).

This was achieved by hybridising 200pmol of each oligo in 50ul of 10xSSC at room temperature for 1 hour. Each well was then washed twice with 100ul of 10xSSC. The Fluorescein labelled RNA produced from the 700 and 800bp fragments were then denatured by heated at 65°C for 5 minutes. One sixth of each of the RNAs, in 50ul of 10xSSC, were then added to a well containing cap 35B and cap 53B. The RNAs were then incubated on ice for 30 minutes to allow the RNA to hybridise to the oligos. The wells were then washed 3 times with 100ul of 10xSSC and the fluorescence measured with a Biolumin 960 (Molecular Dynamics) using Xperiment 1.1.0 software.

The wells were then washed 3 times with 100ul of 5xSSC, 3 times with 100ul of 0.1xSSC and the fluorescence measured as before

- 17 -

after the 3 washes with 5xSSC and 0.1xSSC.

Readings are in Relative Fluorescent Units with background subtraction.

cap 35B:

	10xSSC	5xSSC	0.1xSSC
800bp fragment	66974	220	0
700bp fragment	44350	133	0

cap 53B:

	10xSSC	5xSSC	0.1xSSC
800bp fragment	50734	5136	686
700bp fragment	42660	1438	217

The above data demonstrate that the cap53B oligo immobilised to the microtitre plate can selectively capture RNA produced from the T3 RNA promoter site of pBluescript SK (+). This data also demonstrates that the RNA can be washed of with high stringency washes (e.g. 0.1xSSC).

*Capture of RNA by cap 53B and Probing with 4mers*

100ul of 15xSSC containing 200pmol of cap 53B oligo was transferred to each of 64 wells of a black, streptavidin coated microtitre plate (Boehringer Mannheim) and incubated at room temperature for 2 hours.

The wells were then washed twice with 150ul of 15xSSC. RNA transcribed from the 700 and 800bp clone in a 200ul reaction (see above) was denatured for 2 minutes at 85°C. Both the 700 and 800bp RNAs (see above for production) were then each sub-divided between 24 wells in 100ul of 15xSSC with 16 units of ribonuclease

- 18 -

inhibitor (Boehringer Mannheim) and left to hybridise on ice for 1 hour. The wells were then washed twice with 150ul of ice cold 15xSSC. 200pmol of each 4mer in 100ul of 15xSSC was added to 4 wells (3 with RNA and one with just cap 53B) for both species of RNA and left to hybridise for 1 hour on ice. Each well was then washed 3 times with 150ul of ice cold 15xSSC and the fluorescence read as before (see above). The wells were washed again with 150ul of ice cold 15xSSC and the fluorescence measured.

Sequence of 4mers (all Fluorescein labelled) TACT, TACA, TAAA, TGTC, GTCA, ACCA, TGAA, CAGG.

All data points are expressed as Relative Fluorescent Units are the mean of the three wells minus the reading obtained from the well without RNA.

For the 800bp fragment:

4-mer	1st reading	2nd reading
TACT	0	655
TACA	2292	562
TAAA	0	0
TGTC	1118	382
GTCA	498	57
ACCA	0	0
TGAA	0	0
CAGG	0	116

For the 700bp fragment:

4-mer	1st reading	2nd reading
TACT	756	135
TACA	0	0
TAAA	875	259
TGTC	658	0

- 19 -

GTCA	0	0
ACCA	0	0
TGAA	536	156
CAGG	0	0

The above data demonstrate that the 4mers TACA, TGTC and GTCA hybridise to the RNA transcribed from the 800bp fragment in descending order of strength. These oligos correspond to the following numbered positions in the primary sequence using the 'A' of the start codon as base 1 in TNFa.

For the 800bp fragment:

TACA    264-267  
         276-279  
         392-395

TGTC    60-63  
         256-259  
         770-773

GTCA    255-258  
         615-618

There is an obvious clustering of all three oligos, which bind, around bases 250 to 280.

Oligos found in the sequence but did not bind

TAAA    751-754  
ACCA    371-374  
         374-373  
         568-571  
CAGG    114-117  
         147-150

262-265  
336-339  
390-393  
527-530  
566-569  
696-699  
737-740

For the 700bp fragment:

The oligos TAAA, TACA and TGAA hybridise to the RNA transcribed from the 700bp fragment in descending order of strength. These oligos correspond to the following numbered positions in the primary sequence using the 'A' of the start codon as base 1 in TNFa.

TAAA    838-841  
         1146-1149  
         1168-1171  
         1175-1178  
         1179-1182  
         1186-1189  
         1190-1193  
         1194-1197  
         1210-1213  
         1351 1356

TGAA    873-876  
         921-924  
         1461-1464

TACA is found within the T3 RNA promoter site of pBluescript SK(+).

Oligos found in the sequence but did bind

TACA    1206-1209

- 21 -

	1246-1249
	1313-1316
	1449-1452
TGTC	970-973
	1051-1054
	1105-1108
	1267-1270
GTCA	969-972
	1050-1053
	1398-1401
ACCA	1001-1004
	1395-1398
CAGG	866-869
	888-891
	967-970
	1232-1235
	1321-1324

The fact that different sets of 4mers hybridise to the 2 different RNA species provides evidence that the signal obtained is the product of true hybridisation and not due to a non-specific binding property of the particular 4mers. This is backed up further by the fact that a 4mer with a 100% AT content, and hence lowest  $T_m$ , has hybridised while a 4mer with a 75%GC content, and therefore a much greater  $T_m$ , has not hybridised.

Also, all the oligo sequences are found in both RNA species but only a selection of the oligos were able to bind, and in one instance clustered around a particular sequence. This clustering of 4-mers suggests that this section of the RNA is single stranded or is accessible and is, therefore, an obvious target for antisense.

- 22 -

Therefore, this experiment provides evidence that the single strand parts of an RNA molecule can be probed, and sequence information obtained, by the hybridisation of a 4mer.



CLAIMS:

1. A method for identifying an antisense oligonucleotide capable of binding to a target mRNA, which comprises contacting the target mRNA with each member of an oligonucleotide library separately under hybridisation conditions, optionally removing unhybridised material, and determining which member or members hybridise; wherein the oligonucleotide library comprises a plurality of distinct nucleotide sequences of a predetermined common length, and wherein each nucleotide sequence comprises a known sequence of 4 to 8 bases and all possible combinations of the known sequence are present in the library.
2. A method according to claim 1, wherein each member of the oligonucleotide library is contacted with the target mRNA in a separate container.
3. A method according to claim 2, wherein the target mRNA is immobilized in the container.
4. A method according to claim 1, wherein each member of the oligonucleotide library is immobilized at a separate location on a hybridisation array.
5. A method according to any one of the preceding claims, wherein the length of the known sequence is from 4 to 6.
6. A method according to any one of the preceding claims, wherein each member of the oligonucleotide library comprises a set of nucleotide sequences, each nucleotide sequence of which having a window region comprising the known sequence and a flanking region of no more than 8 bases, wherein all possible combinations of bases in the flanking region are present in each set and the common length of the nucleotide sequences is no more than 12 bases.
7. A method according to any one of the preceding claims, wherein the common length of the nucleotide sequences is no more

than 10 bases.

8. A method according to any one of the preceding claims, wherein the step of determining which member or members hybridise comprises determining the member or members which hybridise more rapidly.

9. A method according to any one of the preceding claims, wherein the sequence of each member which hybridises is compared with sequence information relating to the target mRNA so as to identify in the target mRNA one or more antisense binding sites.

10. A method according to any one of the preceding claims, wherein the nucleotide sequences are made from DNA analogues.

11. A method according to any one of the preceding claims, wherein the target mRNA or the members of the oligonucleotide library are labelled with a label and hybridisation is determined by detection of the label.

12. A method according to claim 11, wherein the label comprises a fluorescent label or a mass label.

13. A method according to claim 11, wherein a detection surface is used to detect the proximity of the label thereto.

14. A method according to claim 13, wherein the label comprises a radiolabel and the detection surface comprises a scintillant.

15. Use of an oligonucleotide library in a method for identifying an antisense oligonucleotide capable of binding to a target mRNA, which library comprises a plurality of distinct nucleotide sequences of a predetermined common length wherein each nucleotide sequence comprises a known sequence of 4 to 8 bases, all possible combinations of the known sequence are present in the library, and each member of the library is

- 25 -

spatially separate from one another.

16. Use according to claim 15, wherein each member of the oligonucleotide library is present in a separate container.

17. Use according to claim 15, wherein each member of the oligonucleotide library is located separately on a hybridisation array.

18. Use according to any one of claims 15 to 17, wherein the length of the known sequence is from 4 to 6.

19. Use according to any one of claims 15 to 18, wherein each member of the oligonucleotide library comprises a set of nucleotide sequences, each nucleotide sequence of which having a window region comprising the known sequence and a flanking region of no more than 8 bases, wherein all possible combinations of bases in the flanking region are present in each set and the common length of the nucleotide sequences is no more than 12 bases.

20. Use according to any one of claims 15 to 19, wherein the common length of the nucleotide sequences is no more than 10 bases.

21. Use according to any one of claims 15 to 20, wherein the nucleotide sequences are made from DNA analogues.

22. Use according to anyone of claims 15 to 21, wherein the target mRNA or the members of the oligonucleotide library are labelled with a label and hybridisation is determined by detection of the label.

23. Use according to claim 22, wherein the label comprises a fluorescent label or a mass label.

# INTERNATIONAL SEARCH REPORT

Inte Application No  
PCT/GB 97/02722

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 C12Q1/68 C12N15/11 C07H21/00 A61K31/70

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 95 21265 A (ISIS INNOVATION ;SOUTHERN EDWIN MELLOR (GB); MIR KALIM ULLAH (GB);) 10 August 1995	1-7, 9-12, 15-23
Y	see the whole document	8,13,14
Y	EP 0 614 988 A (HOECHST AG) 14 September 1994 see the whole document	8
Y	WO 91 08489 A (PACKARD INSTRUMENT CO INC) 13 June 1991 see the whole document	13,14
	--- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

24 February 1998

Date of mailing of the international search report

02/03/1998

Name and mailing address of the ISA  
European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Molina Galan, E

# INTERNATIO SEARCH REPORT

Int. Application No  
PCT/GB'97/02722

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MISHRA R K ET AL: "IN VITRO SELECTION OF ANTISENSE OLIGONUCLEOTIDES TARGETED TO A HAIRPIN STRUCTURE" LIFE SCIENCES, vol. 317, November 1994, pages 977-982, XP002006445 cited in the application ---	
A	CLOAD S T ET AL: "SELECTION OF STRUCTURE-SPECIFIC INHIBITORS OF THE HIV REV-REV RESPONSE ELEMENT COMPLEX" JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 116, no. 2, 26 January 1994, pages 437-442, XP000570388 ---	
A	HO S P ET AL: "POTENT ANTISENSE OLIGONUCLEOTIDES TO THE HUMAN MULTIDRUG RESISTANCE -1 MRNA ARE RATIONALLY SELECTED BY MAPPING RNA-ACCESSIBLE SITES WITH OLIGONUCLEOTIDE LIBRARIES" NUCLEIC ACIDS RESEARCH, vol. 24, no. 10, 15 May 1996, pages 1901-1907, XP002006548 ---	
P,X	WO 97 10332 A (BRAX GENOMICS LTD ;SCHMIDT GUNTER (GB)) 20 March 1997 see the whole document -----	1,5-7,9, 10

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. Application No

PCT/GB 97/02722

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9521265 A	10-08-95	EP 0742837 A JP 9508791 T	20-11-96 09-09-97
EP 0614988 A	14-09-94	CA 2116305 A FI 940830 A JP 6261798 A NO 940608 A	25-08-94 25-08-94 20-09-94 25-08-94
WO 9108489 A	13-06-91	EP 0502123 A	09-09-92
WO 9710332 A	20-03-97	AU 6939496 A	01-04-97